

**SVEUČILIŠTE U ZAGREBU**  
**PRIRODOSLOVNO-MATEMATIČKI FAKULTET**  
**BIOLOŠKI ODSJEK**

**EVOLUCIJA ALTERNATIVNOG PREKRAJANJA RNA**

**EVOLUTION OF ALTERNATIVE RNA-SPLICING**

**SEMINARSKI RAD**

Bruno Pavletić

Preddiplomski studij molekularne biologije

(Undergraduate Study of Molecular Biology)

Mentor: doc. dr. sc. Nenad Malenica

Zagreb, 2018.

# SADRŽAJ

1. UVOD: ZAŠTO PROUČAVATI EVOLUCIJU ALTERNATIVNOG PREKRAJANJA?..	1
2. MEHANIZAM (ALTERNATIVNOG) PREKRAJANJA mRNA .....	2
2.1. Tipovi alternativnog prekrojanja .....	8
3. RAZVOJ INTRONA KAO GLAVNIH META PREKRAJANJA .....	10
3.1. Evolucija i porijeklo introna.....	11
3.1.1. Introni grupe I i II .....	11
3.1.2. tRNA-introni.....	13
3.1.3. Spliceosomalni introni .....	13
4. EVOLUCIJA I PORIJEKLO SPLICEOSOMA .....	15
4.1. Glavne funkcionalne podjedinice spliceosoma .....	17
4.1.1. snRNA .....	17
4.1.2. Sm i Lsm obitelji proteina .....	20
4.1.3. Regulatorni proteini prekrojanja .....	22
5. ZAKLJUČAK: EVOLUCIJSKA VAŽNOST ALTERNATIVNOG PREKRAJANJA.....	24
6. LITERATURA .....	25
7. SAŽETAK .....	28
8. SUMMARY .....	28

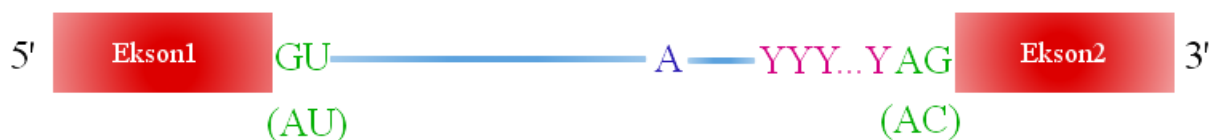
# 1. UVOD: ZAŠTO PROUČAVATI EVOLUCIJU ALTERNATIVNOG PREKRAJANJA?

Ulaskom u eru sekvenciranja genoma, znanstvenici su dobili mogućnost proučavati živi svijet na posve drukčiji način: kroz informaciju kojom su njegovi procesi kodirani. Sve bržim i jeftinijim sekvenciranjem različitih genoma i njihovim usporedbama dobivamo detaljan uvid u razvoj života na Zemlji. Od početka je bilo očekivano da su eukariotski genomi puno veći od prokariotskih tj. da kodiraju za više informacija zbog mnogo veće kompleksnosti građe eukariotskih stanica i organizama. Takva očekivanja su i ostvarena, ali međusobnom usporedbom genoma eukariota uočeno je da veličina genoma nije proporcionalna kompleksnosti organizma (Cooper, 2000). Na to najbolje ukazuje organizam s najvećim poznatim genomom, ameba *Polychaos dubium*, jednostanični organizam s genomom procijenjene veličine 670 Gpb (Friz, 1968) što je oko 210 puta više od ljudskog genoma, veličine oko 3.2 Gpb. Najveće iznenađenje uslijedilo je kada je iz slijeda ljudskog genoma bioinformatičkim alatima dobiveno samo oko 19 000 gena-kandidata koji kodiraju proteine i oko 2000 nekodirajućih RNA gena (Ezkurdia i sur., 2014). Postavljeno je pitanje kako možemo od samo ~20 000 gena dobiti tako kompleksan organizam kao što je čovjek. Dio odgovora na to pitanje daje nam postojanje alternativnog prekrajanja transkripata većine gena koji kodiraju proteine. Prekrojanje ili *splicing* je post-transkripcijska modifikacija RNA kojom se izrezuju nekodirajući dijelovi primarnog transkripta. Takvih nekodirajućih sljedova ima iznimno mnogo u genima velike većine eukariota. Alternativno prekrojanje predstavlja mogućnost stanice da kontrolira koje će genske varijante eksprimirati u ovisnosti o okolišnim ili razvojnim čimbenicima. Time može dobiti različite proteinske produkte. To ukazuje da je proces alternativnog prekrojanja izvor velike funkcionalne raznolikosti i zbog toga je od velikog fundamentalnog i aplikativnog značaja proučavanje porijekla i razvoja ovog sustava.

## 2. MEHANIZAM (ALTERNATIVNOG) PREKRAJANJA mRNA

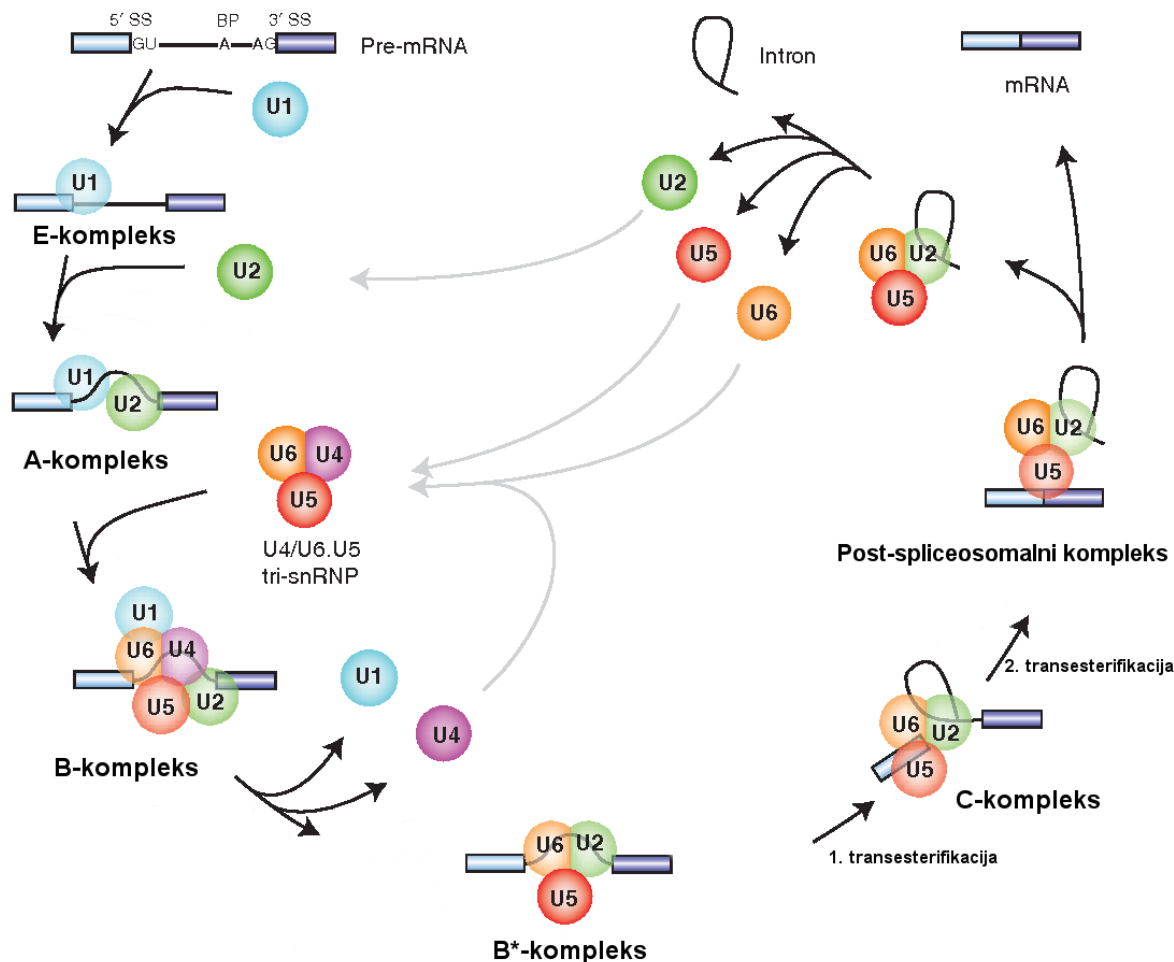
Prije razumijevanja evolucije alternativnog prekrajanja, važno je upoznati se s mehanizmom njegovog djelovanja. Prekrojanje mRNA podrazumijeva izrezivanje nekodirajućih ili nefunkcionalnih dijelova (introna) iz primarnog transkripta kako bi se dobio zreli transkript. Kodirajući dijelovi mRNA koji ostaju u zrelom transkriptu nazivaju se eksoni. Iako postoje četiri tipa introna, od kojih se svaki prekroja na jedinstven način (o čemu će biti riječ kasnije), samo jedan tip introna uključuje alternativno prekrojanje. Za sam proces prekrojanja odgovoran je spliceosom – ribonukleoproteinski kompleks karakterističan za eukariote čija je jedina funkcija prekrojanje nezrele mRNA (pre-mRNA) u kontinuirani kodirajući slijed. U izgradnji spliceosoma sudjeluje velik broj različitih proteina i RNA molekula, a njegovo djelovanje podložno je preciznoj regulaciji drugih faktora što i omogućuje alternativno prekrojanje. Uspješno izrezivanje introna spliceosomom ovisi o nekoliko faktora: o proteinima koji se vežu oko mjesta koja će se prekrojati i tako dovode spliceosom na pravu lokaciju; o razini kondenzacije kromatina u području koji se prepisuje; i o samoj brzini prepisivanja mRNA. Potrebno je naglasiti da spliceosom ne djeluje na gotov transkript, već kotranskripcijski-za vrijeme nastajanja mRNA.

Prekrojanje započinje od pre-mRNA koja sadrži intron. Kako bi ga spliceosom prepoznao, intron sadrži očuvane sljedove koji ga otkrivaju. To su GU (rjeđe AU) slijed na 5' mjestu prekrojanja (5' *splice site*, 5' SS), AG (rjeđe AC) slijed na 3' mjestu prekrojanja (3' SS), *branching* adenin (*branching adenine*, bA) i polipiramidinski slijed (Roy i Irimia, 2014; Slika 1).



**Slika 1.** Struktura klasičnog eukariotskog introna u pre-mRNA. Zeleno su označena mjesta prekrojanja - 5' SS s lijeve strane i 3' SS s desne. U zagradama su sljedovi koji se rjeđe pojavljuju na tim mjestima. Ružičasto je označen polipirimidinski slijed varijabilne duljine bogat citozinima i uracilima (označeni s Y). Tamno plavo je označen *branching* adenin važan u reakciji prekrojanja. Svjetlo plavim linijama su označeni slabo očuvani sljedovi. Intron je omeđen s dva eksona koja se spoje nakon prekrojanja.

U prekrajanju sudjeluje pet malih jezgrinih ribonukleoproteina (snRNP: U1, U2, U4, U5 i U6). Svi ostali faktori su proteinski. Slika 2 prikazuje korake u prekrajanju RNA. Prvi korak je asocijacija U1 snRNP-a na 5' SS lokus čime nastaje E-kompleks, te nakon toga dva faktora, *U2 auxiliary factor-35* (U2AF35) i *U2 auxiliary factor-65* (U2AF65) za 3' SS. Nakon toga, protein *branchpoint binding protein* (BBP) veže se za bA (*branching adenine*). Takva struktura dovodi 5' SS i 3' SS introna u neposrednu blizinu. Potom U2 zamjenjuje BBP na bA mjestu čime nastaje tzv. A-kompleks. Na A-kompleks se za 5' SS, uz već vezani U1, veže se i U4/U6·U5 tri-snRNP te time nastaje tzv. B-kompleks. U4 ima funkciju regulatora prekrajanja jer se preko svoje snRNA komplementarno sparuje s U6 snRNA i time inhibira njegovo djelovanje. Treba doći do disocijacije U4 od U6 kako bi se U6 snRNA sparila s U2 snRNA i započela reakciju prekrajanja, a za što je potrebna hidroliza ATP-a. Hidrolizu obavlja protein Prp8 bez kojeg spliceosom ne bi mogao mijenjati svoju konformaciju i tako postao nefunkcionalan. Nastankom B-kompleksa, dolazi do strukturnih promjena zbog kojih se događa disocijacija U1 i U4 i asocijacija U2 i U6. Pored toga, dolazi do vezanja kompleksa Prp19/CDC5L na mjesto U1 pa se B-kompleks katalitički aktivira (tzv. B\*-kompleks, Fica i sur., 2013). B\*-kompleks katalizira prvu reakciju transesterifikacije što rezultira nastankom tzv. C-kompleksa koji potom katalizira drugu reakciju transesterifikacije potpuno izrezujući intron. Time nastaje post-spliceosomalni kompleks koji se sastoji od izrezanog introna vezanog uz U2/U6·U5 tri-snRNP i zrele mRNA (Slika 2).



**Slika 2.** Shematski prikaz svih koraka prekrajanja RNA. Prvo nastaje E-kompleks tako što se U1 snRNP vezže se za pre-mRNA nakon čega dolazi U2 snRNP (prizvan proteinima U2AF35 i 65 te BBP) pa time nastaje A-kompleks. On na sebe veže U4/U6·U5 tri-snRNP čime se formira B-kompleks. Ta se struktura katalitički aktivira nakon disocijacije U1 i U4 snRNP i spajanja U2 i U6 snRNA te katalizira prvu reakciju transesterifikacije dajući C-kompleks. On potom katalizira drugu reakciju transesterifikacije nakon čega nastaje post-spliceosomalni kompleks od kojeg disocira U2/U6·U5 tri-snRNP.

(modificirano od Will i Lührmann, 2011)

Važno je naglasiti da su 3' SS i 5' SS mjesta iznimno male specifičnosti (samo po dva nukleotida na početku i kraju introna) te da će u transkriptima postojati veliki broj „lažnih“ mjesta prekrajanja. Zato u intronima i eksonima postoje sljedovi na koje se vežu faktori prekrajanja i pomažu definirati ekson ili intron. Hoće li se neki intronski slijed izrezati, u potpunosti ovisi o komunikaciji između faktora prekrajanja. Temelj za mogućnost alternativnog prekrajanja leži u sposobnosti spomenutih faktora da pronađu mjesta koja se trebaju izrezati. Postoje dva osnovna načina kako se ti faktori mogu međusobno prepoznati

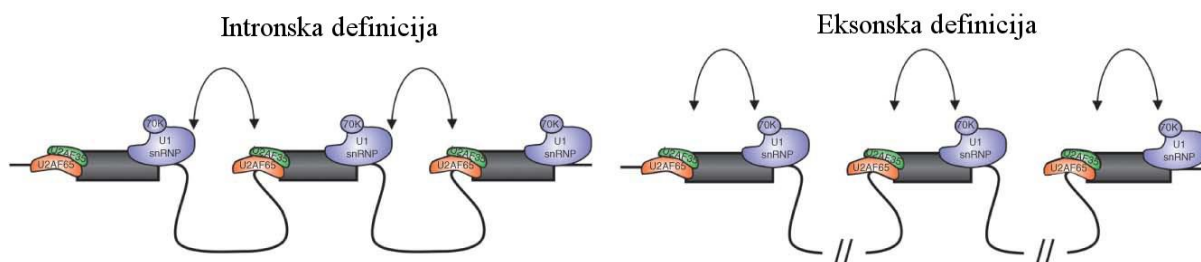
omeđujući slijed koji će biti izrezan iz pre-mRNA: intronska definicija (*intron definition*) i eksonska definicija (*exon definition*, Robberson i sur., 1990). Oko mjesta prekrajanja postoje regulatorni slijedovi koji predstavljaju signal za prekrajanja. Ako regulatorni slijed ima velik afinitet za proteinski faktor prekrajanja tj. predstavlja jaki signal za prekrajanje, susjedni ekson će uvijek ostati dio mRNA pa kažemo da je takav ekson konstitutivan. Ako regulatorni slijed ima slabi afinitet za faktor prekrajanja tj. predstavlja slabi signal prekrajanja, obično dolazi do tzv. *exon skippinga* odnosno izrezivanja (preskakanja) tog eksona zajedno s okolna dva introna. Takav ekson zovemo alternativni. Proteini koji pojačavaju ili oslabljuju signale za prekrajanje mogu biti tkivno-specifični čime kontroliraju u kojem tkivu će prevladavati koja varijanta mRNA.

Eksonska definicija podrazumijeva da faktori prekrajanja interagiraju preko eksona tj. kada U1 pronade 5' SS, on na sebe veže U2AF35/65 na 3' SS prethodnog introna. To znači da se u ovom mehanizmu faktori prekrajanja poslože na kraj jednog introna i početak drugog tako da se između njih nalazi ekson koji će biti uključen u zreli transkript (Slika 3). Ovaj mehanizam pronalaženja mjesta prekrajanja zahtjeva malu udaljenost između dva kraja eksona. Zbog toga je karakterističan za kratke eksone omeđene dugim intronima što je slučaj kod većine genoma sisavaca. Uzevši u obzir da vrijedi ovaj model pronalaska mjesta prekrajanja, postavlja se pitanje kako se izrezuju prvi i zadnji intron u pre-mRNA? Naime, 5' SS prvog introna nema uzvodno 3' SS na koje će se vezati U2AF35/65 i vezati U1. Zbog toga, U1 na prvom intronu interagira s proteinskim kompleksom vezanim za 5'-kapu pre-mRNA. Stoga je jedna od uloga U1 i da štiti RNA od razgradnje i pomaže u njezinom prenošenju kroz jezgrinu membranu. Na isti način, zadnji intron nema nizvodni 5' SS pa zbog toga ni U1 s kojim bi se U2AF35 mogao vezati i tako pronaći zadnji 3' SS. To zadnje mjesto prekrajanja U2AF35 uspješno pronalazi jer umjesto s U1, interagira s proteinima vezanim za 3'-poli(A) rep koji također doprinosi stabilnosti mRNA i omogućuje njen transport u citoplazmu. Valja naglasiti da je ovo samo oblik komunikacije između dva faktora s ciljem pronalaska krajeva introna, no ne ujedno i njihova međusobna interakcija pri katalizi. Izrezivanje slijeda između dva eksona uvijek će se vršiti „preko tog slijeda“ tj. njegov 5' kraj će se približiti 3' kraju i doći će do izrezivanja. Na taj način, eksoni se nikada neće izrezati s posljedicom spajanja dva introna, ali mogu pomoći proteinima i RNP-ovima da pronađu mjesta prekrajanja. To čine specifičnim slijedovima koje sadrže, na koje se vežu proteini koji reguliraju izrezivanje. Takvi proteini mogu biti pojačivači eksonske definicije pa će intron oko kojeg su eksoni sa slijedovima na koje se oni vežu biti češće izrezivan. Također, to mogu biti i utišivači koji umanjuju vjerojatnost izrezivanja introna

pa se takvi regulatori obično vežu oko alternativnih eksona. Evolucija regulatora prekrajanja povezana je uz evoluciju eksonskih sljedova na koje se vežu. Takvi sljedovi obično se nalaze na krajevima eksona. Budući da se proteini regulatori vežu za specifične sljedove eksona, domene koje vežu takve sljedove nemaju veliku mogućnost divergencije jer reguliraju izrezivanje mnogih eksona u genomu. S druge strane, sljedovi eksona na koje se vežu takvi regulatorni proteini mogu divergirati do određene mjere jer će na taj način utjecati samo na učestalost izrezivanja svojih okolnih introna. Budući da je takva promjena kvantitativna, najčešće nije podložna snažnom selektivnom pritisku. Također, eksonski sljedovi imaju veću mogućnost divergencije zbog redundancije genetičkog koda. Iako eksoni kodiraju informaciju, više različitih sljedova može dati isti produkt dok je kod proteina najčešće važno da aminokiselinski slijed bude minimalno izmijenjen jer su takve promijene podložne selekciji.

Mehanizam intronske definicije podrazumijeva da faktori prekrajanja međusobno interagiraju preko introna odnosno faktori prekrajanja se lokaliziraju na kraj jednog eksona i početak drugog eksona tako da se između njih nalazi intron koji će biti izrezan iz pre-mRNA (Slika 3). Taj model komunikacije je učestaliji kod vrsta s velikim eksonima i kratkim intronima. To su obično vrste koje su prošle kroz smanjenje količine introna (Lang i Spritz, 1983) na razini cijelog genoma npr. poput dva modelna organizma *Saccharomyces cerevisiae* (Goguel i Rosbash, 1993) i *Drosophila melanogaster* (Hawkins, 1988). Mehanizam smanjenja se bazira na tome da introni u sebi sadrže sljedove koji ih definiraju i privlače regulatore faktora prekrajanja. Kod ovog fenomena očekivali bi nešto veću konzerviranost takvih intronskih sljedova, nego kod ostatka istog introna i onih sljedova koji se izrezuju pomoću eksonske definicije. To je zato što, za razliku od eksona, intron nema pripadajući proteinski produkt i slobodan je snažno divergirati jer na njega ne djeluje selekcija, već samo genetički drift. Učinak drifta se može zanemariti jer djeluje samo na male populacije i nema utjecaj kada se sljedovi uspoređuju između vrsta, rodova, porodica i ostalih filogenetskih kategorija tj. čak i ako intron ostane konzerviran unutar populacije, ipak će između sve viših filogenetskih razina sve značajnije divergirati.





**Slika 3.** Shematski prikaz intronske i eksonske definicije kod odabira sljedova koji će se izrezati iz pre-mRNA. Kod intronske definicije faktori prekrajanja sastavljaju na dva eksona tako da okružuju intron koji će biti izrezan, a kod eksonske definicije na dva introna oko eksona koji će biti uključen u zrelu mRNA.

(modificirano od De Conti i sur., 2013)

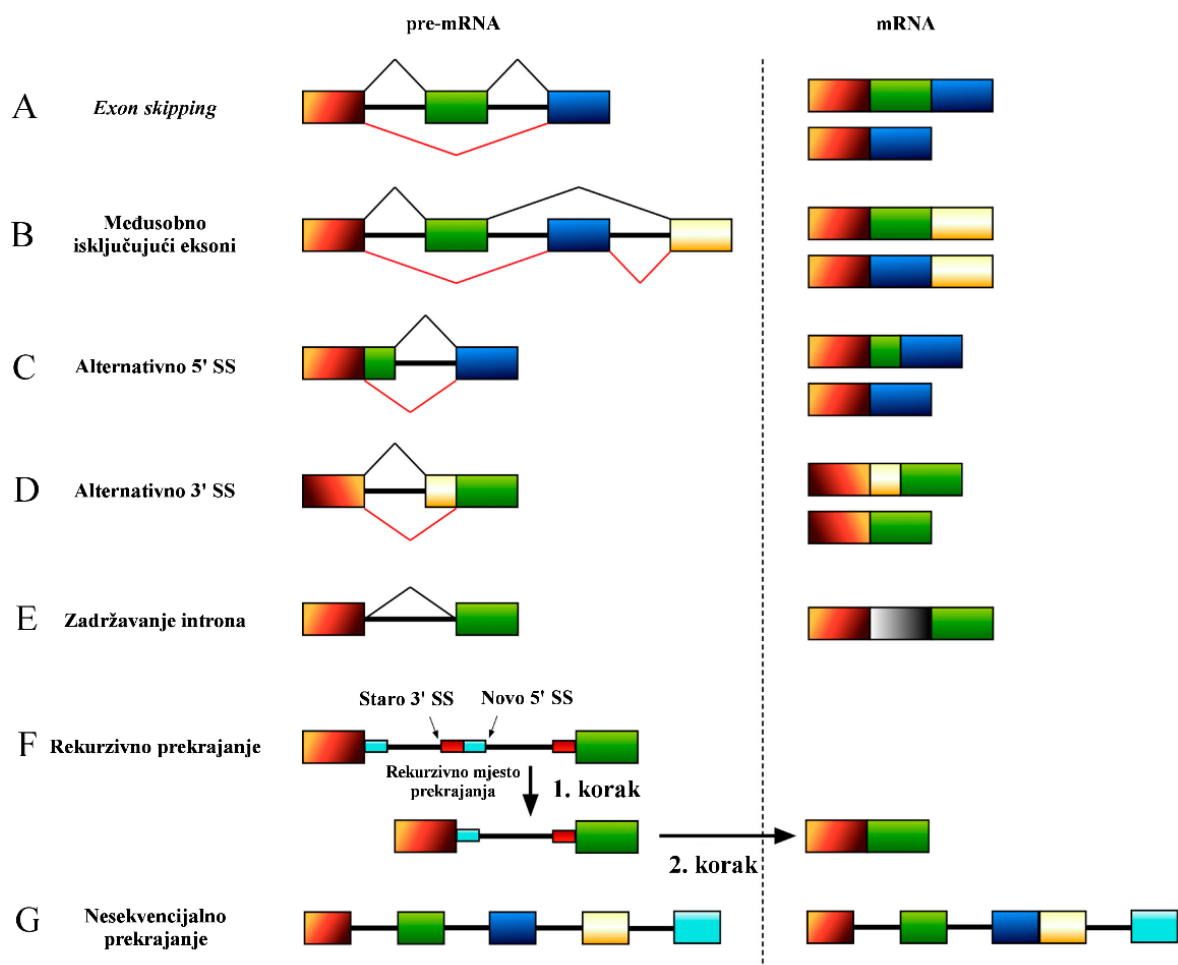
Kod organizama koji većinom koriste eksonsku definiciju za identifikaciju mjesta prekrajanja, očekivano je da će se prilikom alternativnog prekrajanja kratki eksoni omeđeni dugim intronima biti rjeđe izrezani od ostalih. S druge strane, kod organizama kod kojih prevladava intronska definicija očekujemo upravo suprotni efekt: kratki eksoni omeđeni dugim intronima biti će češće izrezani iz pre-mRNA od ostalih. Takvi se trendovi i primjećuju u organizmima (Fox-Walsh i sur., 2005) što je snažan indikator koevolucije duljine introna i eksona s mehanizmima pronalaska mjesta prekrajanja tj. trans-djelujućim regulatornim proteinima i cis-djelujućim regulatornim sljedovima za koje se vežu. Ako u jednom trenutku u evoluciji prevlada eksonska definicija prekrajanja, intronsko-eksonska arhitektura genoma teži smanjenju konstitutivnih eksona i povećanju introna kao kod kralježnjaka, a ako prevlada intronska definicija, teži povećanju konstitutivnih eksona i smanjenju introna kao kod beskralježnjaka i jednostaničnih eukariota. Važno je imati na umu da to vodi u mogućnost regulacije alternativnog prekrajanja neovisno o pojačivačima i utišivačima prekrajanja. U genomima čiji eksoni općenito teže smanjivanju, dugi eksoni češće će se alternativno izrezivati, nego kraći. Također, vrijedi i obrnuto pa je način pronalaska mjesta prekrajanja važan evolucijski faktor.

## 2.1. Tipovi alternativnog prekrajanja

Da bi razumjeli kako je alternativno prekrajanje povezano sa strukturom proteina, važno je imati na umu da u većini eukariotskih genoma jedan ekson najčešće kodira otprilike jednu domenu proteina (Liu i sur., 2005). Alternativno prekrajanje može se obaviti na više različitih načina (Slika 4). Jedan način je, kao što je prethodno opisano, *exon skipping* (Slika 4, A) koji nastaje izrezivanjem eksona zajedno s dva susjedna introna. To je najjednostavniji način za izmjenu cijele jedne domene proteina ili barem njezinog velikog dijela. To je ujedno i najčešći mehanizam alternativnog prekrajanja kod životinja što ukazuje na njegovu važnost u stvaranju velike raznolikosti različitih staničnih linija. Također, postoji mehanizam međusobno isključujućih eksona (Slika 4, B) gdje je izrezivanje dva obično susjedna eksona regulirano na način da izrezivanje jednog onemogućuje izrezivanje drugog eksona. Ovakvi eksoni obično kodiraju istu domenu proteina, ali različite aminokiselinske sljedove pa su korisni u izmjeni proteinske aktivnosti bez velikog utjecaja na tercijarnu, a time i potencijalno kvartarnu strukturu (npr. ako je protein podložan regulaciji drugim proteinom koji prepoznaje njegovu strukturu). Nadalje, u organizmima su također bitni tzv. alternativni 5' SS i 3' SS sljedovi (Slika 4, C i D) koji omogućuju alternativno izrezivanje samo jednog djela eksona – kod alternativnog 3' SS izrezuje se 5' kraj eksona (kod proteina odgovara dijelu bliže N-terminalu), a kod alternativnog 5' SS se izrezuje 3' kraj eksona (kod proteina odgovara djelu bližem C-terminalu). Najčešće podrijetlo ovog oblika alternativnog prekrajanja je kada nastane pseudo-mjesto prekrajanja koje se zadrži u populaciji jer predstavlja neutralnu mutaciju ili donosi selektivnu prednost.

Uz opisane mehanizme, također je moguće da RNA može zadržati cijeli intron tj. preskočiti oba mjesta prekrajanja (Slika 4, E). To je najčešći oblik alternativnog prekrajanja kod biljaka obzirom da one imaju mnogo kraće introne od životinja te uključivanje introna u strukturu proteina nema veliki učinak. Takav način prekrajanja daje intronu selektivni značaj jer mu se jako smanji mogućnost nakupljanja mutacija obzirom da mora kodirati informaciju tako da ne naruši strukturu proteina čiji je dio postao. Također postoje i iznimno rijetki mehanizmi prekrajanja samo kod nekih vrsta: rekurzivno (Slika 4, F) i nesekvencijalno (Slika 4, G) prekrajanje. Za rekurzivno prekrajanje karakteristične su dvije reakcije izrezivanja (četiri reakcije transesterifikacije). Ono je posljedica direktnog ponavljanja i velike duljine introna. Iz toga proizlazi da najčešće nastaje duplikacijom intronskog slijeda pri čemu se stari 3' SS nađe blizu novog 5' SS te to čini novo, kriptičko mjesto prekrajanja. Taj dio introna se izrezuje u

prvom koraku i u drugom se dogodi prekrajanje između starog 5' SS i novog 3' SS. Naravno, ovaj mehanizam nastaje postepeno akumulacijom nukleotida između starog 3' SS i novog 5' SS jer su na početku preblizu da bi ih spliceosom prepoznao. Drugi rijedak način prekrajanja je nesekvencijalno prekrajanje kod kojeg se introni ne izrežu redosljedom karakterističnim za ostale introne, tako da se oni s 5' strane izrežu ranije, već se neki introni koji su s 3' strane izrežu prije njih onih s 5' strane.

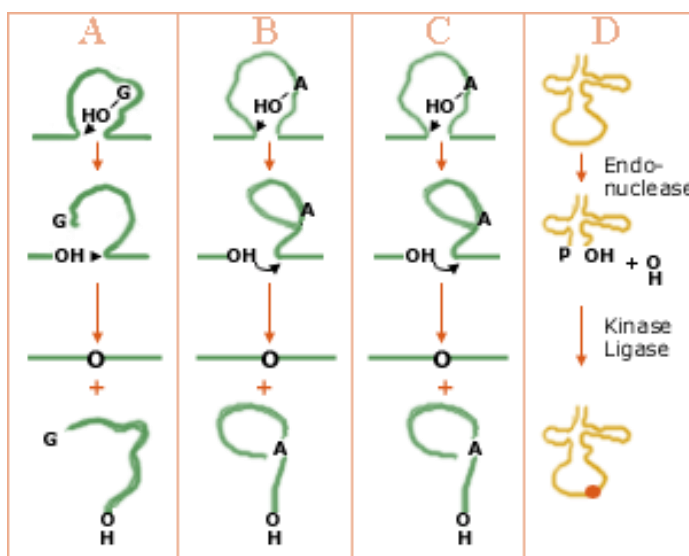


**Slika 4.** Alternativni načini prekrajanja pre-mRNA dokazani u eukariotima. Svaki je način označen posebnim slovom. Dok je *exon skipping* najčešći oblik alternativnog prekrajanja kod životinja, zadržavanje introna je daleko najčešća kod biljaka. Ostali oblici su nešto češći kod životinja, ali postoje mehanizmi karakteristični samo za rijetke, duge transkripte nekih životinja koji prolaze kroz rekurzivno ili nesekvencijalno prekrajanje. Crvenim linijama omeđeni su alternativno izrezani sljedovi.

(Modificirano od Porazinsky i Ladomery, 2018)

### 3. RAZVOJ INTRONA KAO GLAVNIH META PREKRAJANJA

Velik dio genoma ne kodira proteine, a takvi sljedovi mogu biti intergenski ili intragenski. Obje vrste sljedova služe regulaciji staničnih procesa, ali za alternativno prekrajanje RNA najvažniji su intragenski nekodirajući sljedovi tj. introni. Oni postoje u genomima svih triju velikih domena života gdje se nalaze između kodirajućih dijelova gena (eksona). Introni se prepisuju u početni transkript (pre-mRNA) ali budući da ne kodiraju informaciju za konačni produkt, moraju se izrezati kako bi nastao funkcionalan protein. Postoje četiri različite grupe introna (Slika 5): grupa I, grupa II, tRNA introni i spliceosomalni introni. Slika 5 prikazuje sva četiri mehanizma prekrajanja RNA. U ovom radu najveća pažnja posvećuje se spliceosomalnim intronima jer su oni jedini koji prolaze kroz alternativno prekrajanje.



**Slika 5.** Usporedba mehanizma izrezivanja četiri poznate skupine introna. Introni grupe I (A) izrezuju se u potpunosti autokatalitički i njihovo izrezivanje ovisi o prisutnosti gvanina, ili gvanozin trifosfata u okolini. Introni grupe II (B) i spliceosomalni introni (C) izrezuju se istim mehanizmom – tako da njihov adeniozin započne transesterifikaciju. Dok se introni grupe II mogu izrezivati autokatalitički, spliceosomalni introni trebaju pomoć proteina u procesu. Introni vezani uz tRNA (D) jedini su u kojima RNA ne katalizira reakcije prekrajanja već samo proteini i vjerojatno su imali neovisno podrijetlo.

(modificirano od <https://www.nobelprize.org/educational/medicine/dna/a/splicing/introns.html>)

### 3.1. Evolucija i porijeklo introna

Porijeklo introna je jedan od velikih nepoznanica biološke evolucije iz nekoliko razloga. Jedan je to što predstavljaju genetički materijal koji nema funkciju, a za njihovu sintezu je potrebno mnogo energije pa bi bilo logično očekivati da će selekcijski pritisak eliminirati stanice koje prvi puta razviju takve sljedove (Jo i Choi, 2015). To je kontradiktorno zatečenom stanju: mnogi organizmi u prirodi (uključujući i čovjeka) imaju iznimno gusto raspoređene introne po većini gena. Postavlja se pitanje kako su se i zašto introni uspjeli održati u organizmima. Druga nepoznanica je uočena velika razlika u prisutnosti introna kod prokariota i eukariota. Prokarioti općenito imaju jako malo introna dok su među eukariotima oni rasprostranjeni u gotovo svim poznatim genomima. Poznat je samo jedan genom bez introna, i to organelni - nukleomorf alge *Hemiselmis andersenii* (Lane i sur., 2007). To ukazuje na radikalnu divergenciju prekrajanja RNA između prokariota i eukariota. Postavlja se pitanje je li zajednički predak ove dvije skupine imao introne pa su ih prokarioti tijekom evolucije izgubili zbog selektivnog pritiska (hipoteza *introns early*, Darnell, 1978) ili su introni apomorfija (svojstvo koje nije prisutno kod predaka) eukariota (hipoteza *introns late*, Stoltzfus, 1994). Nadalje, sve bržim objavljivanjem sljedova genoma, postaje jasno da gustoća introna jako varira ne samo među vrstama, već i između kromosoma istog genoma (Vaulot i sur., 2012). Zbog takvih rezultata, iznimno je teško postaviti model evolucije introna budući da kod dvije filogenetski udaljene vrste oni mogu biti sličnih duljina i genomske distribucije.

#### 3.1.1. Introni grupe I i II

Introni grupe I i II najstarija su vrsta introna. Oni su najjednostavniji i mogu se izrezati iz RNA autokatalitički tj. bez sudjelovanja drugih makromolekula u dvije reakcije transesterifikacije prikazane na Slici 5. Razlika između ove dvije skupine je u mehanizmu izrezivanja. Introni grupe I u RNA formiraju sekundarnu strukturu koja veže gvanozin (G) ili gvanozin trifosfat (GTP) iz okoline čijom 3'-OH skuponom nukleofilno napada fosfat na svom 5' SS u prvoj reakciji transesterifikacije. To stvara slobodni 3'-OH kraj eksona koji zatim nukleofilno napada prvi fosfat nizvodno od 3' SS introna u drugoj reakciji transesterifikacije (Cech i Bass, 1986).

Introni grupe II, ovisno o klasi, u prvoj reakciji nukleofilno napadaju svoj 5' SS slijed molekulom vode (Podar i sur., 1998) ili 2'-OH evolucijski visoko očuvanog *branching* adenzina (bA) kojeg sadrže u sebi (Peebles i sur., 1986). Nakon toga, slobodni 3'-OH kraj eksona napada prvi fosfat iza 3' SS i izrezuje intron iz transkripta. Kod onih koji napadaju 5' SS očuvanim bA, formira se struktura zvana lariat (omča, „laso“)– RNA molekula koja je svojim 5' fosfatom spojena za svoj vlastiti bA preko njegovog 2' kisika. Ista struktura javlja se kao nusprodukt i onih introna izrezanih spliceosomom, putem istih biokemijskih koraka. Zato je opće prihvaćena hipoteza da su spliceosomalni introni nastali upravo iz introna grupe II koji koriste bA u reakciji izrezivanja (Haugen i sur., 2005). Iako ove dvije skupine introna imaju sposobnost samoizrezivanja, oni *in vivo* rijetko funkcioniraju sami. Najčešće se kataliza odvija unutar kompleksa RNP-a, koji uključuje RNA koja se prekraja, i maturazu –intronom kodiran holoenzim . Ti proteini pomažu ne samo u izrezivanju, već i u kontroli i transpoziciji takvih introna što daje značajnu prednost u odnosu na introne koji se sami izrezuju. Naime, neke maturaze djeluju kao reverzne transkriptaze jer imaju RT-domenu i kataliziraju prijenos izrezanih introna na drugo mjesto u genomu što navodi na zaključak o transpozonskom porijeklu introna.

Nedavno je pokazana i iznimno zanimljiva povezanost introna grupe II sa spliceosomom. Naime, Zhao i Pyle (2016) opisali su kristalnu strukturu RT domene maturaze i napravili filogenetsku analizu tog proteina. Njihovi rezultati ukazuju da su najbliži homolozi RT domene maturaze dva enzima: Prp8 i RNA-ovisna RNA polimeraza flavivirusa. Prp8 je iznimno bitna ATP-aza spliceosoma (spomenuta u prvom dijelu ovog rada) koja mu omogućuje katalizu pa je ovaj rezultat čvrst indikator hipoteze o porijeklu spliceosomalnih introna od onih iz grupe II. Homologija RT domene maturaze s flavivirusnom RNA polimerazom daje indicaciju o porijeklu reverzne transkriptaze (koja je DNA polimeraza) iz RNA polimeraza.

### **3.1.2. tRNA-introni**

Za razliku od ostalih vrsta introna, tRNA-introni arheja i eukariota ne izrezuju se transesterifikacijskim reakcijama, već ih na krajevima izrezuje visoko očuvana obitelj endonukleaza specifičnih za tRNA (InterPro oznaka obitelji: IPR006676), a lijepljenjem zaostalih slobodnih krajeva ligazom nastaje zrela tRNA (Slika 5). Kod bakterija koje imaju introne u genima za tRNA, oni se izrezuju autokatalitički i taj mehanizam najvjerojatnije nije homologan onom kod arheja i eukariota. Primarna struktura endonukleaza koje kataliziraju ovu reakciju visoko je očuvana između eukariota i arheja, što snažno ide u prilog hipotezi o zajedničkom porijeklu mehanizma prekrajanja tRNA. Također, činjenica da se tRNA-introni izrezuju u potpunosti drugačijim mehanizmom od introna grupe I i II snažno navodi na zaključak da su ovi introni nastali neovisno o ostalima tj. nemaju evolucijsku povezanost sa grupom I i II, a time ni s spliceosomalnim intronima. Treba naglasiti da je mogućnost izrezivanja tRNA introna snažan selektivni faktor budući da interferiraju s genima koji utječu na stabilnost cijelog sustava - ako se tRNA ne može vezati za ribosom, aminokiselina se ne može ugraditi u niti jedan protein i negativna selekcija snažno djeluje protiv takve stanice. Zbog toga je većina mutacija koje narušavaju izrezivanje introna iz tRNA letalna.

### **3.1.3. Spliceosomalni introni**

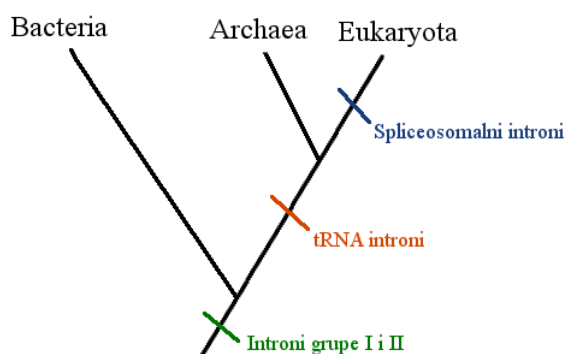
Spliceosomalni introni karakteristični su za eukariote i prekranje ovim mehanizmom je najučestaliji način prekrajanja mRNA kod te skupine. Oni najvjerojatnije imaju zajedničkog pretka s intronima grupe II na što ukazuje identičan mehanizam prekrajanja. Najveća je razlika u tome što se introni grupe II mogu sami izrezivati (iako rijetko u živim stanicama), a izrezivanje spliceosomom u potpunosti ovisi o ribonukleoproteinima. Također, prekranje RNA spliceosomalnim intronima u potpunosti je ovisno o proteinima koji prepoznaju eksone ili introne i dovode spliceosom do mjesta prekrajanja. Zbog toga je prekranje mRNA kod eukariota izrazito podložno regulaciji što predstavlja izvor velike raznolikosti. Regulacijom prekrajanja mRNA omogućeno je dobivanje više različitih proteinskih produkata koji mogu s vremenom razviti različite funkcije u stanici.

Budući da se spliceosomalni introni ne mogu sami izrezivati, više nisu pod selektivnim pritiskom da održe sekundarnu strukturu RNA (važnu za samo-izrezivanje) pa su njihovi

sljedovi mogli jako divergirati tijekom evolucije. Očuvali su se jedino sljedovi koji omogućuju spliceosomu da prepozna i izreže intron (Slika 1). Većina eukariotskih introna su GT-AG tipa tj. na krajevima imaju GT i AG, no postoji mali broj introna u većini poznatih genoma koji su tipa AT-AC. Također, postoje dvije varijante spliceosoma koji izrezuju introne (o čemu će biti riječi u trećem dijelu). Obje ove varijante spliceosoma prepoznaju oba tipa introna kao svoj supstrat (Burge i sur., 1998) što ukazuje na postojanje potencijalne kompeticije između te dvije vrste introna ili kompeticije obje varijante spliceosoma. Možda su negdje u ranoj evoluciji eukariota GT-AG introni potisnuli AT-AC introne koji su se izgubili. Tome u prilog ide i zapažanje da tijekom evolucije introni mogu preći iz AT-AC u GT-AG tip, ali ne i obrnuto (Burge i sur., 1998).

Iako introni ne kodiraju niti proteine niti funkcionalnu RNA, oni ipak u sebi nose informacije koje ih čine funkcionalnim evolucijskim jedinicama. Svaki slijed DNA koji teži tome da se širi populacijom, bez obzira je li koristan za organizam, opstaje u populaciji i u njoj dalje evoluira tj. divergira tijekom vremena. Dok populacijama i vrstama bazu za divergenciju predstavljaju nove ekološke niše, DNA sljedovima to predstavljaju same populacije. Konkretno, što je populacija veća, rasprostranjenija i u slabijem međusobnom kontaktu, veća je divergencija između njenih dijelova. Na isti način, što je neki DNA slijed rasprostranjeniji u populaciji, veće su šanse za mutacije i zbog toga brža divergencija (pogotovo kada za širenje nije toliko važna sama informacija u njoj). U tome vjerojatno leži objašnjenje velike raznolikosti introna. Od trenutka kada su se pojavili, introni su se vjerojatno transpozonski širili drevnim genomima. Najjednostavnija pretpostavka je da im je jedna od prvih bioloških funkcija bila regulacija procesa translacije ili funkcije proteina, što je ovisilo jedino o tome hoće li se intron izrezati iz transkripta ili neće. One eukariotske stanice koje su mogle preciznije regulirati aktivnost više svojih proteina bolje su preživljavale u istom mikrokolišu od onih koje su imale slabu regulaciju. Uz to, otvorila se mogućnost kolonizacije novih mikookoliša. Zbog toga su se vjerojatno introni proširili u mnoge gene ranih eukariota, kao i u nove okoliše (davali su stanicama selektivnu prednost) pa je došlo do snažne divergencije intronskih sljedova. Slika 6 prikazuje najvjerojatniji tijek nastanka sva četiri mehanizma prekrajanja RNA.





**Slika 6.** Pretpostavljeni evolucijski tijek nastanka četiri skupine intronskih sljedova. Bazira se na činjenici da sve domene života sadrže introne grupe I i II iz čega se zaključuje da su te dvije skupine najstarije. Introni specifični za tRNA izrezuju se potpuno drugačijim mehanizmom od svih ostalih introna i smatra se da su neovisno nastali kod pretka eukariota i arheja jer ih te dvije skupine dijele. Spliceosomalni introni specifični su za eukariote, a izrezuju se sličnim mehanizmom kao introni grupe II pa se smatra da su nastali iz njih kod zadnjeg pretka eukariota.

## 4. EVOLUCIJA I PORIJEKLO SPLICEOSOMA

U eukariotskim stanicama, prekrajanje pre-mRNA obavlja veliki kompleks proteina i RNA – spliceosom. On ima nekoliko karakteristika koje ga razlikuju od mnogih drugih proteinskih kompleksa, a čine sličnim ribosomu. Za početak, riječ je o kompleksnoj strukturi. Spliceosom čini više od 100 različitih proteina (kod čovjeka) od kojih većina ima strukturnu ili regulatornu ulogu. Drugo, struktura cijelog kompleksa vrlo je dinamična. Spliceosom, baš kao ni ribosom, ne postoji potpuno formiran u stanici i „čeka“ svoju pre-mRNA kako bi obavio izrezivanje introna. Umjesto toga, njegove komponente se sukcesivno slažu oko pre-mRNA i isto tako disociraju tijekom cijelog procesa. Zbog toga se teško može definirati jedna konstantna struktura za spliceosom. Treće, iako je većinom građen od proteina, reakcija izrezivanja koju obavlja u potpunosti katalizira snRNA. Ove karakteristike ukazuju na dugu evoluciju i drevno porijeklo spliceosoma.

Uz glavni katalitički spliceosomski kompleks koji se sastoji od pet snRNP-a o kojima je bilo riječi u drugom djelu, postoji još mnogo proteina regulatora koji variraju od vrste do vrste. Uz opisani U2 tip spliceosoma, postoji i manji, U12 tip koji koristi drukčije snRNA (Tablica 1). Oba tipa spliceosoma koriste U5 snRNA, najvjerojatnije jer ona ne sudjeluje u

prepoznavanju intronskih mjesta. Budući da U12 spliceosom koristi drugačije snRNA za navođenje, postoji razred pripadajućih U12 introna koje se više preferira. Ostali introni spadaju u razred U2. Taj razred introna puno je češći u eukariotskim genomima (kod čovjeka 99.5% svih introna). Neke vrste su čak u potpunosti izgubile U12 introne pa im uslijed toga U12 spliceosom, zbog divergencije, više nije ni funkcionalan (Lin i sur., 2010). Pokazano je da tijekom evolucije intron može promijeniti razred u oba smjera ( $U2 \leftrightarrow U12$ ) iako je prijelaz iz U2 u U12 jako rijedak. Štoviše, češće dolazi do gubitka U12 introna, nego takvog prijelaza  $U2 \rightarrow U12$ , što rezultira trendom gubitka U12 introna u genomima (Lin i sur., 2010).

**Tablica 1.** snRNA koje koristi U2 tip spliceosoma i njihovi analozi u U12 tipu. Jedina RNA koju koriste oba tipa je U5.

snRNA U2 spliceosoma	Analogna snRNA u U12 spliceosomu
U1	U11
U2	U12
U4	U4atac
U5	U5
U6	U6atac

Budući da je jedina poznata uloga spliceosoma u svim organizmima izrezivanje introna iz pre-mRNA, pretpostavlja se da je to bila njegova početna uloga. Istraživanja ortolognih gena koji kodiraju za podjedinice spliceosoma ukazuju da je velika većina proteina koji ga danas grade već bila prisutna u zajedničkom pretku svih eukariota (Collins i Penny, 2005). To navodi na zaključak da je takva struktura nastala jednom i vrlo rano, mnogo prije pretka svih eukariota i od tada mu se evolucija prilično usporila. Moguće objašnjenje je da je u jednom trenutku razvoja, kada je već bio jako kompleksan, spliceosom prošao kroz snažnu selekciju efektom uskog grla (*bottleneck effect*) zbog koje je samo jedna verzija opstala (ona kod pretka eukariota). Takav je efekt najvjerojatniji ako je cijela evolucija spliceosoma bila iznimno lokalna. Manje je vjerojatno, ali moguće da se spliceosom proširio po Zemlji, ali je ograničio organizme koji su ga imali u trenutku kada je nastupila nekakva promjena i uzrokovala masovno „izumiranje“ organizama sa svim spliceosomalnim mehanizmima osim jednog. Također je moguće da je razvoj ranog spliceosoma od početka bio postupan tj. često su nastajali proteini koji su ga nadograđivali i uvijek bi značili jaku selektivnu prednost nad njegovom

„prethodnom verzijom“. Taj scenarij također traži međusobnu kompeticiju među ancestralnim jedinkama sa spliceosomom i zahtjeva da se one nalaze u istoj ekološkoj niši relativno uske rasprostranjenosti.

#### ***4.1. Glavne funkcionalne podjedinice spliceosoma***

Kao što je prethodno spomenuto, spliceosom se sastoji od pet centralnih katalitičkih ribonukleoproteinskih kompleksa (U1, U2, U4, U5 i U6) i ostalih proteina koji variraju od vrste do vrste. Proteini najvažniji za prekrajanje RNA osim središnjih RNP-a nose naziv Prp (*pre-mRNA processing*). Velika većina tih proteina ili prepoznaje RNA sljedove ili funkcioniraju kao nukleotidtrifosfataze (NTP-aze) i omogućuju dinamične promjene strukture spliceosoma. Uz Prp, postoje mnoge druge proteinske obitelji koje sudjeluju u funkciji i regulaciji spliceosoma od kojih su za alternativno prekrajanje najvažnije hnRNP (ribonukleoproteini vezani uz heterogenu nuklearnu RNA) i SR (sadrže domenu bogatu ponavljanjima dipeptida serin-arginin tj. **S-R**).

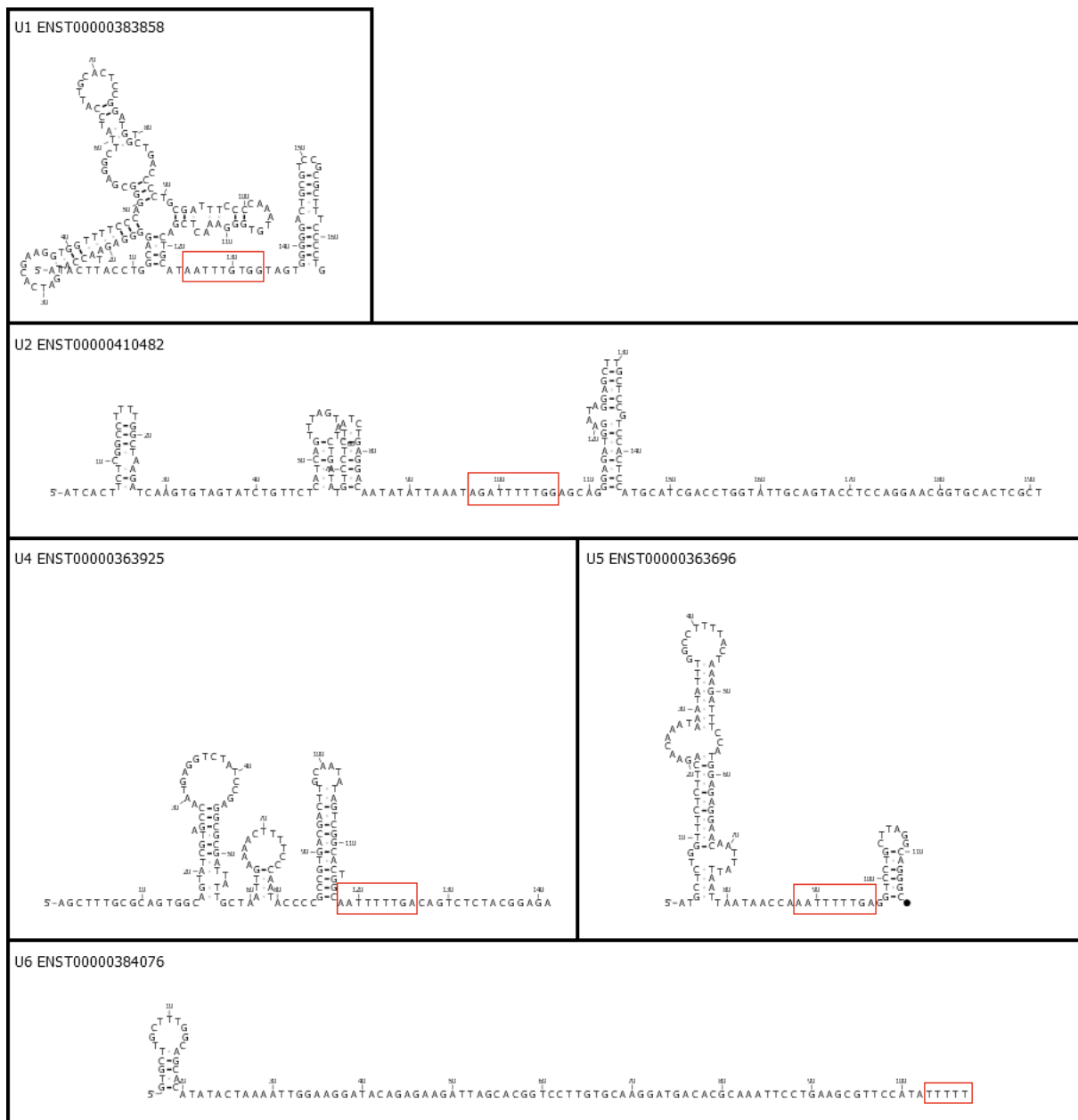
##### ***4.1.1. snRNA***

Katalitički kompleks spliceosoma koristi različite RNA molekule kao sonde za prepoznavanje intronskih sljedova. One spadaju u skupinu snRNA (često se nazivaju U-RNA jer su bogate uridinom). snRNA imaju različite funkcije u stanici, sve vezane uz procesiranje RNA: procesiranje (obrada) ribosomalne RNA (U3, U8-10, U13 i U14; Sloan i sur., 2017), procesiranje 3' kraja histonske mRNA (U7; Chen i Wagner, 2010) i procesiranje pre-mRNA (Tablica 1). snRNA koje sudjeluju u prekrajanju kod metazoa nalaze se u mnogo kopija od kojih su neke aktivni geni, a neki pseudogeni. Unutar vrste sve kopije su skoro identične, ali između vrsta se mogu značajno razlikovati (Jia i sur., 2012) što ukazuje na visoku razinu funkcionalne redundancije i važnost ovih snRNA – fenomen koji se naziva usklađena evolucija. Neki od njih se pojavljuju u skupinama zajedno, a neki su raspršeni po genomu.

U1 snRNA se komplementarno sparuje s 5' mjestom prekrajanja introna čime pripadajući snRNP-1 započinje reakciju. Pretpostavlja se da je ova snRNA nastala iz jednog od tRNA gena koji se nalazio u ponavljajućim kopijama u genomu pretka eukariota. Takav se zaključak izvodi zbog sličnosti u primarnoj i sekundarnoj strukturi U1 snRNA s tRNA genima

(Slika 7; Hogeweg i Konings, 1985). U2 snRNA sparuje se sa slijedom oko bA, ali ga na to mjesto mora dovesti BBP. Ne zna se puno o evoluciji ove snRNA osim da pripadajući mutanti imaju poteškoće u alternativnom prekrajanju općenito, pri čemu je najznačajniji nedostatak izrezivanja introna (Jia i sur., 2012). Zbog toga je vjerojatno i njegova evolucija usko povezana uz mehanizme i regulaciju alternativnog prekrajanja. Evolucija U4 snRNA direktno je vezana uz evoluciju U6 snRNA jer U4 svojim sparivanjem za U6 inhibira njegovu aktivnost. Točna funkcija U5 snRNA je najmanje poznata, ali se pretpostavlja da sudjeluje u povezivanju 5' SS i 3' SS tijekom reakcija transesterifikacije. Zbog nepoznavanja točne funkcije, teško je govoriti o modelu evolucije ove RNA. U6 snRNA je među najbolje očuvanim između vrsta. Zanimljivo je da poliuridinski slijed na 3' kraju U6 snRNA (Slika 7), koji je vezno mjesto za protein Lsm (zajedno s proteinom p110 katalizira sparivanje U4/U6), nije kodiran na genomu, već ga naknadno dodaje enzim terminalna uridililtransferaza 1 (TUT1; Yamashita i sur., 2017). To pokazuje da je U6 snRNA mogla naknadno steći svojstvo vezanja za Lsm, vjerojatno tek pojavom TUT1 enzima. To znači da je do tada djelovala sama ili u kompleksu sa Sm proteinima. Tek pojava novog proteina za interakciju s U6 u formi TUT1 dala je selektivnu prednost takvim organizmima. Važno je naglasiti da U6 snRNA katalizira obje reakcije transesterifikacije u spliceosomu (Fica i sur., 2013) tj. iako u prekrajanju sudjeluju ribonukleoproteini, ono je u potpunosti katalizirano s RNA, dok proteini imaju stabilizacijsku i/ili regulatornu ulogu. To ukazuje na drevno porijeklo prekrajanja spliceosomom pri čemu je U6 snRNA vjerojatno najstariji predstavnik.

Sve RNA vezane uz spliceosom imaju visoko očuvane i primarne i sekundane strukture među eukariotima (Shukla i Padgett, 1999). Među njima međusobno je najslićnije mjesto vezanja proteinskog kompleksa Sm (*Smith*, prezime pacijentice kod koje je otkriven): slijed RAU<sub>1-6</sub>GR gdje R označava purin (G ili A), a U<sub>1-6</sub> 1-6 ponavljanja U (Liautard i sur., 1982). Ovaj slijed ostaje jednolančan u sekundarnoj strukturi snRNA i vežu ga Sm proteini (Slika 7). Jedina iznimka je U6 snRNA koja na 3'-kraju ima poliuridinski slijed za koji se ne veže član skupine Sm proteina, već njegov homolog Lsm (*Like Sm*; Achsel i sur., 1999). Nema dokaza da su snRNA međusobno homologne tj. da su njihovi zajednički preci bili paralozi u drevnom genomu pa je vjerojatno da su nastali neovisno jedni o drugima. Sljedovi koji vežu Sm-domenu konvergentno su evoluirali pod selektivnim pritiskom jer su upravo one snRNA koje se najbolje vežu za proteine prekrajanja mogle i najefikasnije vršiti svoju funkciju.

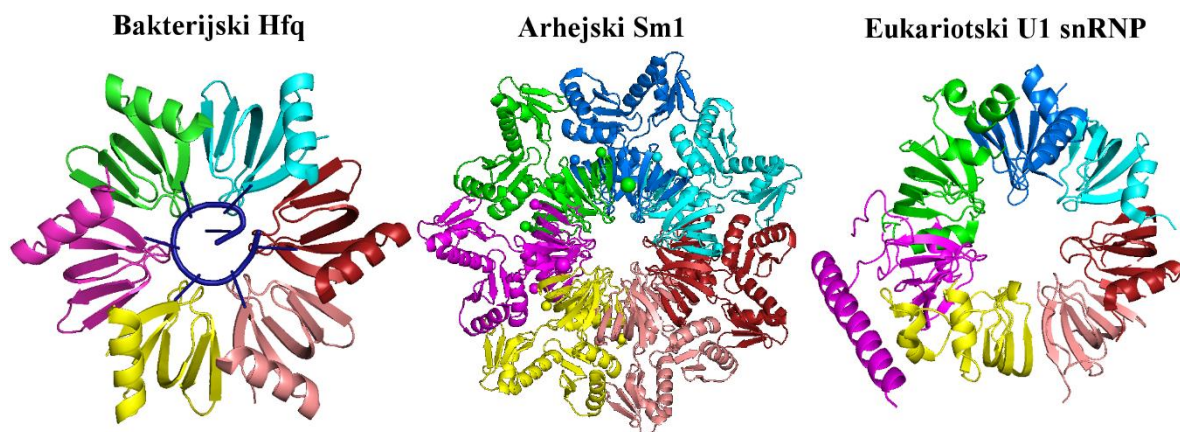


**Slika 7.** Prikaz sekundarnih struktura katalitičkih snRNA spliceosoma čovjeka. Vidljivo je da U1 snRNA ima strukturu „djeteline“ karakterističnu za tRNA pa se pretpostavlja da je iz takve i nastala. Crveno su označeni sljedovi za koje se vežu Sm proteini, osim U6 koji veže Lsm. Iako se radi o RNA sljedovima, ove strukture dobivene su predviđanjem iz sljedova koji se u programe unose u DNA formi (s T umjesto U) pa se na mjestu svakog U nalazi T.

(modificirano iz izvora: <https://www.ensembl.org/>)

#### 4.1.2. Sm i Lsm obitelji proteina

Proteinske komponente ribonukleoproteina spliceosoma spadaju u proteinsku superobitelj LSM. Ona se sastoji od dvije srodne obitelji proteina: Sm i Lsm. Ovo je vrlo široka superobitelj čiji se homolozi pojavljuju u sve tri domene života i funkcije su im u svim organizmima vezane uz funkcioniranje ili biogenezu RNA (He i Parker, 2000; Kufel i sur., 2003). Karakteristika članova LSM superobitelji je sedam visoko očuvanih aminokiselina (tzv. Sm motif) kojima se vežu za ciljnu RNA. Kod bakterija, jedini član Lsm obitelji je protein Hfq (*host factor Q*) koji gradi kružne homoheksamere (Slika 8) i funkcionira kao RNA šaperon za RNA bogatu uridinima ili adeninima (Schumacher i sur., 2002). Kod arheja, ovi proteini nisu dobro istraženi, ali se zna da su dva člana LSM superobitelji koji tvore kružne homoheptamere rašireni i među arhejama – Sm1 i Sm2 (Achsel i sur., 2001). Ti proteini čine funkcionalne homoheptamere (Slika 8). Većina arheja sadrži samo Sm1 dok je kod rijetkih nađen i Sm2. To navodi na zaključak da je Sm1 zasigurno postojao u zajedničkom pretku arheja, a Sm2 je nastao kasnije. Iako su ovi proteini najučestaliji među arhejama, ipak je nađena jedna (*Methanococcus jannaschii*) koja nema niti Sm1 niti Sm2, ali za razliku od svih ostalih ima protein blisko srodan Hfq što navodi na zaključak da se Hfq i Sm proteini mogu komplementirati kod arheja (Nielsen i sur., 2007) zbog vjerojatno sličnih funkcija. Ako je tomu tako, moguće je da je protein sličan Hfq ancestralan svim LSM proteinima pa su iz njega nastali Sm proteini (Sm1) kod pretka arheja i eurariota koji su ga zbog kompeticije za supstrat potisnuli tj. Hfq je mogao nestati kada je Sm1 preuzeo njegovu funkciju. U eukariota je poznato sedam paralognih proteina Sm obitelji: B, D1, D2, D3, E, F i G. Oni zajedno rade kružne heteroheptamere (Slika 8) i tako se vežu za ciljnu RNA molekulu (U1, U2, U5 ili U4/U6) pomoću Sm domene. Ti proteini su dobro očuvani jer čine glavne proteinske komponente spliceosoma. Budući da su eukariotski genomi najčešće bogati intronima, mutacije u Sm proteinima utječu na cijeli transkriptom stanice, a time i proteom. Očito je da bi takva stanica ili organizam teško preživio i dao potomstvo. Zbog toga su eukariotski Sm proteini jako dobro očuvani i primjećuje se njihova ekspanzija od kad su se eukarioti odvojili od arheja, ali i dalje ostaje tendencija da tvore heptamere pa je vjerojatno zbog toga njihova ekspanzija stala. Od proteina Lsm obitelji, nađeno je 8 paraloga u eukariotima (Lsm1-8) od kojih proteini Lsm2-8 čine heteroheptamer koji veže U6, dimer U4/U6 i trimer U4/U6·U5 (Vidal i sur., 1999; Stevens i Abelson, 1999). Lsm proteini uz prekrajanje još sudjeluju i u uklanjanju 5' kape mRNA i time pomažu u njezinoj razgradnji (Lsm1-7, Bouveret i sur., 2000).



**Slika 8.** Strukture proteinskih kompleksa LSM obitelji iz tri domene života. Vidljivo je očuvana prstenasta struktura koja se izgrađuje oko RNA. Radi jednostavnosti, RNA je prikazana samo na bakterijskom kompleksu (plavo) iako svi interagiraju s RNA na sličan način. Kod bakterija, kompleks je izgrađen od šest identičnih podjedinica, kod arheja sedam identičnih podjedinica, a kod eukariota sedam različitih podjedinica. Ovdje su prikazani homoheksamer Hfq iz *Staphylococcus aureus* (PDB ID: 1KQ2), homoheptamer Sm1 iz *Pyrobaculum aerophilum* (PDB ID: 1M5Q) i heteroheptamer U1 snRNP iz *Saccharomyces cerevisiae* (PDB ID: 5UZ5).

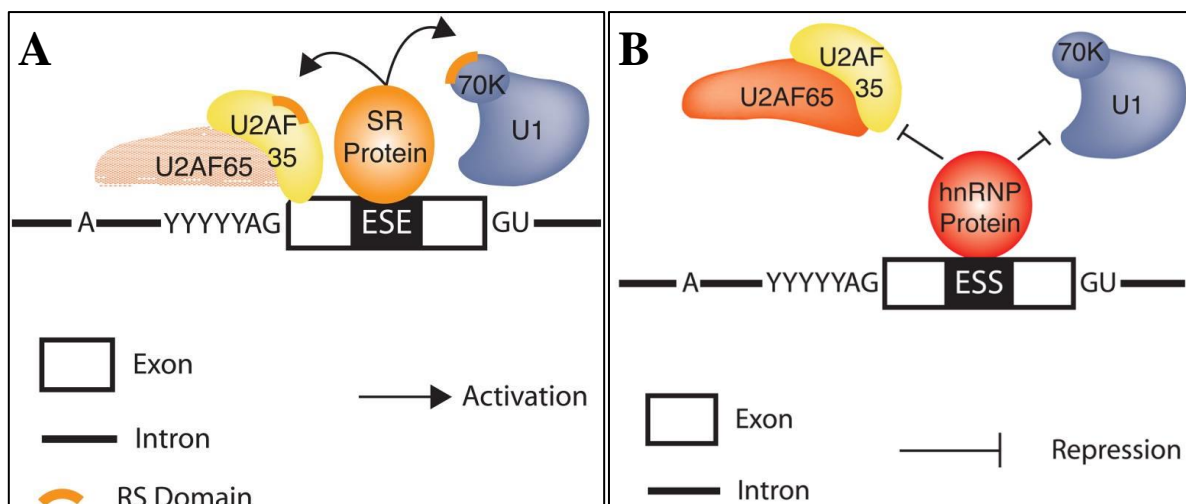
Među proteinima LSM superobitelji vidljiv je trend interakcija s RNA i formiranja prstenastih multimera u svim organizmima u kojima su istraženi. Kod eukariota, ovi su proteini poprimili novu funkciju, a to je prekranje mRNA. To bi mogao biti razlog njihove očite ekspanzije u pretku eukariota. Kao i kod arheja, kompleksi LSM proteina (kod eukariota su dijelovi snRNP-a) formiraju samo heptamere. Dok arheje imaju jedan glavni Sm gen (Sm1) koji čini homoheptamer, eukarioti imaju za svaku podjedinicu heptamera po jedan zasebni gen i čine heteroheptamere. Može se zaključiti da ovi proteini imaju iznimno kompetitivnu evoluciju i kada god se u genomu nađe kodiran neki redundantni protein, oni ga potisnu (za razliku od snRNA koje teže stvaranju svojih redundantnih kopija). To potvrđuje činjenica da kod bakterija i arheja postoji samo jedan rašireni LSM protein, dok ih kod eukariota ima sedam od kojih svi zajedno sudjeluju u prekranju.

#### 4.1.3. Regulatorni proteini prekrajanja

Odabir mjesta prekrajanja ovisi o mnogim faktorima (raspravljeno u prvom dijelu). Jedan od njih su pojačivači prekrajanja koji se nalaze kao DNA sljedovi u eksonima ili intronima. Te sljedove prepoznaju specifični proteini (regulatorni elementi prekrajanja, REP) koji dovode ili uklanjaju komponente spliceosoma i time reguliraju hoće li ekson biti uključen u zreli transkript. Sljedovi na koje se vežu ovi proteini dijele se na četiri skupine: *exon splicing enhancers* (ESE, pojačavaju zadržavanje eksona), *exon splicing silencers* (ESS, smanjuju zadržavanje eksona), *intron splicing enhancers* (ISE, pojačavaju zadržavanje introna) i *intron splicing silencers* (ISS, smanjuju zadržavanje introna). Regulacijom ekspresije REP ili dostupnošću njihovih veznih mjesta, stanice mogu regulirati koje će eksone (ili introne) zadržati, a koje izrezati što je važno za preciznu ekspresiju različitih oblika istog proteina u različitim tkivima posredovanjem alternativnog prekrajanja.

Najbolje okarakterizirane skupine REP-a su SR i hnRNP. Te skupine imaju suprotstavljene funkcije: dok SR proteini vežu ESE i povećavaju vjerojatnost zadržavanja eksona u konačnom transkriptu, hnRNP vežu ESS i povećavaju vjerojatnost izrezivanja (Slika 9). Obje skupine vežu eksonske sljedove i tijekom evolucije su prošle kroz velike ekspanzije duplikacijama i nakupljanjem mutacija (Busch i Hertel, 2012) prema modelu zadržavanja genskih kopija. Prema tom modelu, kada nastane duplikacija gena, jedna kopija ostaje pod selektivnim pritiskom i zadržava originalnu funkciju dok druga može nakupljati mutacije i divergirati dok ne poprimi novu funkciju. Moguće je da je rana ekspanzija ovih proteina dovela do puno veće raznolikosti alternativnog prekrajanja i time do pojave višestaničnosti (Busch i Hertel, 2012). Pojavom mogućnosti izmjene proteinskog produkta u ovisnosti o vanjskim ili unutarnjim signalima zasigurno je potaknula snažnu divergenciju eukariota u njihovim najranijim fazama razvoja.





**Slika 9.** Načini djelovanja SR (A) i hnRNP (B) proteinskih obitelji – dvije skupine proteina najznačajnije za regulaciju prekrajnja RNA.

(izvor: Busch i Hertel, 2012)

Budući da ove skupine reguliraju prekrajnje vezanjem za eksone, oni funkcioniraju prema modelu eksonske definicije. SR proteini dovode faktore prekrajnja interagirajući s njima preko RS domene. Vežu se za ESE i s obje strane eksona mogu dovoditi faktore prekrajnja (Slika 9). S druge strane, hnRNP smanjuju afinitet vezanja faktora prekrajnja za 5' SS i 3' SS i zbog toga se eksoni na koje djeluje preskoče i izrežu s okolnim intronima. Budući da SR proteini nedostaju samo u nekim jednostaničnim eukariotima, opće je prihvaćeno da je ancestralni SR već postojao kod zadnjeg pretka eukariota pa je kod nekih linija izgubljen. Vjeruje se da je širenjem ove skupina proteina, ujedno i povećana mogućnost alternativnog prekrajnja. Takav pojačivač prekrajnja vjerojatno je evoluirao pod evolucijskim pritiskom nastalim mutiranjem sljedova važnih za prepoznavanje samog mjesta prekrajnja. Organizmi koji su imali proteinske pojačivače signala prekrajnja mogli su nakupljati mutacije u tim signalima zbog čega su divergirali. Skupina hnRNP evoluirala je i zbog nastanka pseudo mjesta prekrajnja nasumičnim mutacijama. Mogućnost nakupljanja lažnih mjesta prekrajnja koje stanica može tolerirati daje selektivnu prednost organizmima koji razvijaju takve proteine. U svakom slučaju, SR i hnRNP proteini doprinose velikoj proteinskoj raznolikosti većine eukariota, uključujući i čovjeka. Postoje indicije da su SR i hnRNP blisko srodne skupine tj. da dijele zajedničkog pretka zbog međusobno sličnih domena (Busch i Hertel, 2012). Važno je naglasiti da ove skupine imaju jako povezane funkcije i djeluju na sličan način (posrednici interakcije između RNA i proteina) te je moguće da su njihove sličnosti rezultat konvergentne

evolucije, a ne srodnosti. Na srodnost ovih skupina također ukazuje i činjenica da su obje bile prisutne kod zadnjeg pretka svih eukariota. Dovoljna je samo jedna duplikacija ancestralnog gena za takav scenarij, a u slučaju neovisnog nastanka, te skupine bi trebale prolaziti kroz dugi proces konvergentne evolucije.

## **5. ZAKLJUČAK: EVOLUCIJSKA VAŽNOST ALTERNATIVNOG PREKRAJANJA**

U odnosu na bakterije i arheje, eukarioti pokazuju izrazitu morfološku raznolikost. Takva morfološka raznolikost posljedica je velike raznolikosti transkriptoma, a time i proteoma među eukariotima. Značajna karakteristika koja razlikuje eukariotske transkriptome od ostalih je mogućnost alternativnog prekrajanja RNA. Nije teško vjerovati da je upravo pojava ovog svojstva omogućila snažnu ranu divergenciju eukariotskih linija. Povećanje kodirajućeg potencijala genoma kao posljedica alternativnog prekrajanja RNA omogućilo je stanicama veću fleksibilnost u prilagodbi na okolišne uvjete što je kao posljedicu moglo imati kolonizaciju novih prostora, a time i divergenciju populacija. Selektivni pritisci koji su tijekom evolucije oblikovali regulatore prekrajanja i duljine intorna/eksona djelovale su na razvoj mehanizama alternativnog prekrajanja. Uz već prisutnu mogućnost stvaranja više proteina iz jednog gena, razvoj međustanične komunikacije usklađene s kontrolom ekspresije faktora prekrajanja, omogućio je nastanak prvih višestaničnih organizama s diferenciranim tkivima i različitim tkivno-specifičnih izoformi. Nakon toga, krenula je koevolucija mehanizama prekrajanja i međustanične komunikacije što je dovelo do dodatne divergencije među mnogostaničnim organizmima. Zbog toga, alternativno prekrajanje predstavlja jedan od najvažnijih faktora za divergenciju i razvoj eukariota.

## 6. LITERATURA

Achsel T., Brahms H., Kastner B., Bachi A., Wilm M., Lührmann R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *The EMBO Journal*, 18(20), 5789–5802.

Achsel T, Stark H, Lührmann R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), 3685–3689.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. *The EMBO Journal*, 19(7), 1661–1671.

Burge CB., Padgett RA., Sharp PA. (1998). Evolutionary fates and origins of U12-type introns. *Molecular Cell*, 2(6), 773-785.

Busch, A., & Hertel, K. J. (2012). EVOLUTION OF SR PROTEIN AND HnRNP SPLICING REGULATORY FACTORS. Wiley Interdisciplinary Reviews. *RNA*, 3(1), 1–12.

Cech TR., Bass BL. (1986). Biological catalysis by RNA. *Annual Review of Biochemistry*, 55, 599-629.

Chen, J., Wagner, EJ. (2010). snRNA 3' end formation: the dawn of the Integrator complex. *Biochemical Society Transactions*, 38(4), 1082–1087.

Collins L., Penny D. (2005). Complex spliceosomal organization ancestral to extant eukaryotes. *Molecular Biology and Evolution*. 22(4), 1053-1066.

Cooper, GM. *The Cell: A Molecular Approach*. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. The Complexity of Eukaryotic Genomes.

Darnell JE. Jr. (1978). Implications of RNA-RNA splicing in evolution of eukaryotic cells. *Science*, 202(4374), 1257-1260.

De Conti L., Baralle M., Buratti E. (2013). Exon and intron definition in pre-mRNA splicing. *Wiley Interdisciplinary Reviews. RNA*, 4(1), 49-60.

Ezkurdia, I., Juan, D., Rodriguez, JM., Frankish, A., Diekhans, M., Harrow, J., ... Tress, ML. (2014). Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Human Molecular Genetics*, 23(22), 5866–5878.

Fica, SM., Tuttle, N., Novak, T., Li, N-S., Lu, J., Koodathingal, P., Dai, Q., Staley, JP., Piccirilli, JA. (2013). RNA catalyzes nuclear pre-mRNA splicing. *Nature*, 503(7475), 229–234.

Fox-Walsh, KL., Dou, Y., Lam, B., Hung, S., Baldi, PF., & Hertel, KJ. (2005). The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proceedings of the National Academy of Sciences of the United States of America*, 102(45), 16176–16181.

- Friz, CT. (1968). The biochemical composition of the free-living Amoebae *Chaos chaos*, *Amoeba dubia* and *Amoeba proteus*. *Comparative Biochemistry and Physiology*, 26(1), 81–90.
- Goguel V., Rosbash M. (1993). Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell*, 72(6), 893-901.
- Haugen P., Simon DM., Bhattacharya D. (2005). The natural history of group I introns. *Trends in Genetics*, 21(2), 111-119.
- Hawkins, JD. (1988). A survey on intron and exon lengths. *Nucleic Acids Research*, 16(21), 9893–9908.
- He W., Prker R. (2000). Functions of Lsm proteins in mRNA degradation and splicing. *Current Opinion in Cell Biology*, 12(3), 346-350.
- Hogeweg P., Konings DA. (1985). U1 snRNA: the evolution of its primary and secondary structure. *Journal of Molecular Evolution*, 21(4), 323-333.
- Jia Y., Mu JC., Ackerman SL. (2012). Mutation of a U2 snRNA Gene Causes Global Disruption of Alternative Splicing and Neurodegeneration. *Cell*, 148(1), 296-308.
- Jo, B-S., Choi, SS. (2015). Introns: The Functional Benefits of Introns in Genomes. *Genomics & Informatics*, 13(4), 112–118.
- Kufel J., Allmang C., Petfalski E., Beggs J., Tollervey D. (2003). Lsm Proteins are required for normal processing and stability of ribosomal RNAs. *The Journal of Biological Chemistry*, 278(4), 2147-2156.
- Lane, CE., van den Heuvel, K., Kozera, C., Curtis, BA., Parsons, BJ., Bowman, S., Archibald, JM. (2007). Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), 19908–19913.
- Lang KM., Spritz RA. (1983). RNA splice site selection: evidence for a 5' leads to 3' scanning model. *Science*, 220(4604), 1351-1355.
- Liautard JP., Sri-Widada J., Brunel C., Jeanteur P. (1982). Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. *Journal of Molecular Biology*, 162(3), 623-643.
- Lin, C-F., Mount, SM., Jarmołowski, A., Makałowski, W. (2010). Evolutionary dynamics of U12-type spliceosomal introns. *BMC Evolutionary Biology*, 10, 47.
- Liu, M., Wu S., Walch H., Grigoriev A. (2005). Exon–domain correlation and its corollaries. *Bioinformatics*, 21(15), 3213-3216.
- Nielsen, JS., Bøggild, A., Andersen, CBF., Nielsen, G., Boysen, A., Brodersen, DE., Valentin-Hansen, P. (2007). An Hfq-like protein in archaea: Crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*. *RNA*, 13(12), 2213–2223.

- Podar M., Chu VT., Pyle AM., Perlman PS. (1998). Group II intron splicing in vivo by first-step hydrolysis. *Nature*, 391(6670), 915-918.
- Porazinsky S. i Lodomery M. (2018). Alternative Splicing in the Hippo Pathway—Implications for Disease and Potential Therapeutic Targets. *Genes*, 9(3), 161.
- Robberson, BL., Cote, GJ., & Berget, SM. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and Cellular Biology*, 10(1), 84–94.
- Roy SW., Irimia M. (2014). Diversity and evolution of spliceosomal systems. *Methods in Molecular Biology*, 1126, 13-33.
- Schumacher MA., Pearson RF., Møller T., Valentin-Hansen P., Brennan RG. (2002). Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *The EMBO Journal*, 21(13), 3546-3556.
- Shukla, GC., Padgett, RA. (1999). Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. *RNA*, 5(4), 525–538.
- Sloan, KE., Warda, AS., Sharma, S., Entian, K-D., Lafontaine, DLJ., Bohnsack, MT. (2017). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology*, 14(9), 1138–1152.
- Stevens, SW., Abelson, J. (1999). Purification of the yeast U4/U6-U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 96(13), 7226–7231.
- Stoltzfus A. (1994). Origin of introns--early or late. *Nature*, 369(6481), 526-527.
- Vaulot D., Lepère C., Toulza E., De la Iglesia R., Poulain J., Gaboyer F., ... Piganeau G. (2012). Metagenomes of the picoalga Bathycoccus from the Chile coastal upwelling. *PLoS One*
- Vidal, VP., Verdone, L., Mayes, AE., & Beggs, JD. (1999). Characterization of U6 snRNA-protein interactions. *RNA*, 5(11), 1470–1481.
- Will CL., Lührmann R. (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology*, 3(7), a003707.
- Yamashita S., Takagi Y., Nagaike T., Tomita K. (2017). Crystal structures of U6 snRNA-specific terminal uridylyltransferase. *Nature Communications*, 8, 15788.
- Zhao, C., Pyle, AM. (2016). Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nature Structural & Molecular Biology*, 23(6), 558–565.
- <https://www.ensembl.org/>

## 7. SAŽETAK

Pojava alternativnog prekrajanja nesumnjivo predstavlja iznimno važan korak u evoluciji eukariota. Iako nije poznato kada se točno pojavilo, zasigurno je od tada snažno povećalo kodirajući potencijal genoma. Ono je nastalo kao svojevrsna nadogradnja na prekrajanje pre-mRNA. Prekrajanje RNA nastalo je kao posljedica pojave introna u ranim genomima koji su se morali izrezivati iz transkripata. Introni najvjerojatnije vuku transpozonsko porijeklo što im je omogućilo brzo širenje drevnim genomima. Kompleksnost sustava za prekrajanje samo se povećavala tijekom evolucije, ali osnovni mehanizam ostao je isti. Otprilike u vrijeme odvajanja eukariota od arheja, u ranom pretku eukariotske skupine pojavili su se snRNA i Sm proteini koji su asociirali u ribonukleoproteine (snRNP) koji su katalizirali prekrajanje te je time značajno smanjen selektivni pritisak na intronske sljedove jer mogućnost izrezivanja više nije ovisila o vlastitoj sekundarnoj strukturi. Uz to, snRNP-i su mogli uspostaviti interakciju s novonastalim skupinama proteina – SR i hnRNP koji su ih regulirali. To predstavlja začetak spliceosoma i alternativnog prekrajanja, a obje karakteristike su bile već prisutne kod najkasnijeg zajedničkog pretka eukariota. Od tada, središnje jedinice spliceosoma su do danas ostale jako dobro očuvane među svim proučenim eukariotima, ali je regulacija alternativnog prekrajanja divergirala i izravno utjecala na građu i način života eukariotskih organizama.

## 8. SUMMARY

Appearance of alternative splicing undoubtedly marks a crucial step in eukaryotic evolution. Even though it's not clear when exactly this phenomenon appeared, it must have strongly increased the coding potential of genomes. It clearly appeared as an upgrade to pre-mRNA splicing. RNA splicing came about as a consequence of introns existing in early genomes, which should have been cut out from transcripts. Introns probably derived from transposons, which initially gave them the ability to spread across ancient genomes quickly. The complexity of the splicing system increased through the process of evolution, but the main mechanism remained the same. Approximately at the time of eukaryotic divergence from archaea, in early ancestor of eukaryotic group, snRNA and Sm proteins appeared that associated into ribonucleoproteins (snRNP) and could catalyse RNA splicing. This strongly

relaxed the selection pressure on intronic sequences because their ability to be spliced out didn't depend on their secondary structure anymore. Also, snRNPs were able to interact with newly arisen protein groups – SR and hnRNP which regulated them. This marks the beginning of spliceosome and alternative splicing and both characteristics were already present in the last eukaryotic common ancestor. Since that time, core spliceosomal subunits remained highly conserved between all studied eukaryotes, but alternative splicing regulation diverged and directly influenced the structure and lifestyle of eukaryotic organisms in general.